

January 15, 1962

L-1012

Professor John W. Pratt
Graduate School of Business Administration
Harvard University
Soldiers Field
Boston 63, Massachusetts

Dear John:

I appreciated very much receiving your detailed comment on my paper, "Risk, Ambiguity and the Savage Axioms," and I am embarrassed that it has taken me so long to reply. I had expected long before this to be able to send you a paper embodying some of the work I have done on the subject since writing that paper, which would deal with your criticisms among other problems; but since I have nothing yet in a form to send out, let me comment directly on the point you raise.

Your counterexample is essentially the same one presented by Howard Raiffa in his comment on my paper in the QJE Symposium; I presume you have seen this. Harry Markowitz also raised this example when I discussed my ideas with him; and of course, Howard and I had discussed it long ago in connection with his experiments with his classes. Considerable reflection on the point didn't lead me to change my views then, nor does it now. Let me try to explain why.

Let me take the example discussed by Raiffa in his comment in the QJE (November, 1961, pp. 693-94); the argument is exactly the same as for the one you consider. We imagine an urn containing 30 red balls and 60 black and yellow balls, the latter in unknown proportion of black to yellow. A subject is offered four acts as follows:

		State		
		Red	Black	Yellow
Act	I	\$ 100	0	0
	II	0	\$ 100	0
	III	\$ 100	0	\$ 100
	IV	0	\$ 100	\$ 100

As Raiffa agrees, "most subjects choose act I over II and IV over III, thus violating the Savage Axioms" and "many of these subjects are reluctant to change their choices when it is pointed out to them that their behavior is

inconsistent with the Sure-thing Principle." (p. 294. I would be interested to know your own initial, intuitive reaction to these choices; there are those who are never even tempted to violate the axioms, but I believe that Raiffa is not one. Are you?)

Raiffa proposes--as do you--to "undermine their confidence in their initial choices" by asking them to consult their intuition in the following situation, which I did not discuss. "In option A a fair (unbiased) coin is tossed and act I is taken if heads appears, whereas act IV is taken if tails appears; with option B, heads leads to act II and tails to act III." Thus:

	<u>Heads</u>	<u>Tails</u>
Option A:	Act I	Act IV
Option B:	Act II	Act III

Raiffa continues: "The final outcomes of either option depend on the toss of the coin and the selection of a ball. Now let's do the accounting by analyzing the implications of the options conditional on the color of the withdrawn ball. Our analysis takes the form:

	<u>Red</u>	<u>Black</u>	<u>Yellow</u>
Option A	(An "objective" 50-50 chance of \$100 and \$ 0	Same	Same
Option B	Same	Same	Same

"But this reasoning should lead everyone to assert that options A and B are objectively identical....these options look awfully alike to me!"

Well, they look alike to me, too. Of course the options are identical, for purposes of decision-making; as you put it, "they give identical, 'unambiguous' probability distributions over the consequences." I agree with you thoroughly that it would seem unreasonable to prefer one option to the other. I wouldn't, any more than you would. Where does this apparently unsuspected harmony of minds leave us?

It leaves you and Raiffa with a mistaken prediction on my choice between these two options. But let us examine your reasons for the prediction. I will admit you had plausible grounds (I haven't proved my sanity yet, though I think I'm safe). As Raiffa puts it, "by strict dominance option A is better than B" for me since I prefer I to II and IV to III. Your corresponding argument is: "It certainly does not seem reasonable to me to make a decision you will soon regret (after flipping)" (and in this case, you imply, I would "wish" I had chosen option B, "regretting" my indifference between A and B, after the coin had been tossed; whichever way it fell).

To answer your comment first: if, after flipping I experienced a feeling of "regret" that I had been indifferent between the two options, wishing I had definitely chosen option B, I would struggle to repress this feeling as foolish, unreasonable, and altogether unworthy of me, using precisely your argument that before flipping, a careful analysis showed the two options to be identical for all practical purposes. It would be unfair to myself (and I try never to be unfair to myself) to reproach myself or to suffer regret for being indifferent between two alternatives with identical consequences.

But this leaves Raiffa's argument: given my initial preferences between pairs of actions, option A dominates option B. Yet I say that I am indifferent between option A and option B. "Something must give!" Raiffa exclaims; "on thinking it over, wouldn't you like to change your mind about your initial preferences?"

No.

On thinking it over, I still notice that in comparing action I to action II, the former offers me a definite " $1/3$ chance" of \$100 and a " $2/3$ chance" of \$0, while with the latter my chance of \$100 is something (unknown) between 0 and $2/3$ (while the chance of \$0 is, correspondingly, between $1/3$ and 1). This statement seems to me to have a clear intuitive meaning, but there are various ways I could give it an explicit operational meaning; I mention several below. If I were offered a choice between actions I and II, I would still prefer I. Similarly, in choosing between III and IV, I would still prefer IV; it offers me a definite, $2/3$ chance of \$100, where the chance of \$100 with action III is somewhere between $1/3$ and 1. If asked to explain my principle of decision, I might say: "If two bets differ in the 'vagueness' of their probability distribution over payoffs, I will prefer the one whose distribution of payoffs is more 'definite' unless the 'best guess' distribution of payoffs corresponding to the other is distinctly superior." Since the "best guess" distributions of payoffs corresponding to the "vague" bets II and III are no better than the definite distributions corresponding, respectively, to I and IV, they don't carry the "ambiguity premium" that would make me indifferent between them and the corresponding unambiguous bets.

I can't deny that, with these preferences, option A does dominate option B. What, then, do I say to Raiffa's comment: "I cannot see how anyone could refute the logic leading to the conclusion that given your initial choices, you should prefer option A to option B."? How, that is, could anyone stand up and confess in public that he was willing, in any circumstances, to violate the principle of admissibility: that he was indifferent between two alternatives one of which strictly dominated the other? I agree with Raiffa that "Something must give," and since in my case I can't honestly deny my pattern of preferences, I must assert that for me, in these circumstances, Dominance is not a normative criterion.

I admit that I was surprised the first time I found myself wilfully violating the criterion of domination. I would not have predicted that I would ever behave that way. I will even admit that, while I shed no tears in discarding the Sure-thing Principle (I hadn't lived with it as long as Savage had; easy come, easy go), I felt some pangs when I took the old mottoes like "Never take a dominated action" off my wall. But it's like everything else; I've learned to live with myself again.

For one thing, my loyalty to the notion of dominance hasn't vanished; it has just been qualified. I simply wouldn't apply it automatically now, without checking my actual preferences, when the payoffs or prizes in a set of bets were themselves ambiguous bets (rather than "sure things"). As the above example illustrates, the effect of combining in a single bet, with unambiguous probabilities, two ambiguous bets, can be to produce a bet which has an unambiguous probability distribution of outcomes. The result is a "lottery ticket" which I may prefer to either of its (ambiguous) prizes!

It is easy to miss this essential point in focusing on the choice between option A and option B. It is true that, despite the fact that A dominates B, I am indifferent between A and B; but it is no less significant to note that I prefer option B to both of the prizes it offers, actions II and III considered individually. Without going through the whole argument, I merely note that option B has an unambiguous distribution of payoffs (.5 chance of \$100, .5 chance of \$0) while actions II and III do not (the better of them, action III, has between $1/3$ and 1 probability of \$100).

Thus, the options Raiffa offered his students, in which they were allowed to choose the color on which they would bet in the "ambiguous" urn, were essentially different from those I discussed in my paper. I drew attention to this difference in footnote 7, p. 651. (footnote 7 in P-2173) The way Raiffa poses the problem, I would agree that they "should" regard the various mixed options as identical. But I have a question for Raiffa. Now that we are admitting the possibility of mixed options (having assumed the existence of some event with "unambiguous" probabilities for the subject), suppose we offer his students the following choice: (a) they can bet on Red in the "unknown" urn; (b) they can flip a coin between betting on Red and betting on Black in the "unknown" urn; (c) they can bet on Red in the "known" urn. As I say, I agree with Raiffa that it should be possible to convince them that (b) and (c) are equivalent. But are they indifferent between (a) and (b)? I'll bet they aren't; I'll give odds that a significant number of them prefer (b) to (a); and that violates the axioms as much as any of the examples discussed in my paper! Moreover, I'll bet that after all his argument designed to make them indifferent between (b) and (c), a lot of them will still prefer (c) to (a). This is precisely my pattern of preferences, which we've been discussing above; it carries all the same consequences, including the horrid implication that Raiffa harbors in his statistics classes students who might, under the right circumstances, (a) prefer a dominated action, and (b) prefer a lottery ticket offering prizes A and B with definite probabilities, both to A and to B. (If he flunks them, I will try to get them slots at RAND.)

I promised earlier to elaborate on the operational meanings one might supply for statements like, "my subjective probability for event H is between $1/3$ and $2/3$." This letter is long enough now, and perhaps I should let this ride until I can send you a full exposition of my thoughts on this subject since writing the article. However, following are some comments that may be suggestive. When I say that action I offers me a "definite" probability of $1/3$ for the payoff \$100, one thing I might mean by it is: in a very large number of situations in which my subjective expectation of "winning" were equivalent to the expectation I feel with respect to action I, I would expect to win with a frequency very close to $1/3$ of the times. That is not true of my expectation about the average results of a large number of bets subjectively similar to action II. In a large number of such situations, I would expect with high confidence that the fraction of wins would not exceed $2/3$. But I would not be at all surprised if I won less than, say, $1/6$ of the time and not too surprised if I never won. In this particular interpretation, I can point to the difference between feeling that I know the average results to be expected (to give odds on) in a large series of trials involving the same subjective uncertainties, and feeling that I don't know them, except within wide limits.

This is not the only operational interpretation that can be put on such statements. Others involve betting behavior, (see C. A. B. Smith's article in the Journal of the Royal Statistical Society, Vol. 23, No. 1, January 1961) or comparisons with events of "definite" probability (see Koopman's articles on intuitive probability, e.g., Annals of Mathematics, Vol. 41, No. 2, April, 1940).

In terms of betting behavior: let us suppose that my utility function has been measured (by bets with unambiguous events), and let us assign the values 0, 1 to \$0, \$100 in the example above. Suppose that I were willing to give up to $1/3$ utils for act I, and up to $2/3$ utils for act IV; I would be acting "as if" I assigned $1/3$ probability to Red (and, properly, $2/3$ to not-Red). If, following Smith, we speak of the value to me of act I as the "lower betting probability" for Red and (1-the value of act IV) as the "upper betting probability" for Red, the two would be equal in this case.

But suppose that the most I would give for act II should be $1/6$; and the most I would give for act III would be $1/2$. The "lower betting probability" for Black would then be $1/6$, and the upper betting probability $(1-1/2)=1/2$. If the set of values I assign to various bets is "coherent", the upper betting probability must be greater than or equal to the lower betting probability; statements by de Finetti, Ramsey, and Savage that the requirement of "coherence" implies equality of the upper and lower betting probabilities, and hence the "definiteness" of all probabilities measured by acceptable betting odds, are simply incorrect. (Try "making a book" against me! It can't be done here, where the sum of my upper and lower betting probabilities is less than 1, though it could be done if they added to more than 1).

Or, to follow the Koopman approach, suppose someone asked me to compare my degree of belief that I would draw a black ball from this urn with my degree of belief that I would draw a black ball from an urn containing exactly six balls, n of which were black. I might state that $n=1$ was the largest n such that I could definitely state that the first was not smaller than the second, and that $n=5$ was the smallest n such that I could definitely state that the second was not smaller than the first. For the event "not-Black," the corresponding numbers should be (if my beliefs are "coherent") $n=3$ and $n=5$. Thus, in Koopman's terms, my upper and lower probabilities for Black are $1/2$ and $1/6$.

Suppose that my beliefs concerning the probabilities of Red, Black and Yellow, as shown by the above sorts of statements are represented by: $1/5$, $1/6-1/2$, $1/6-1/2$. Suppose further that my utilities are known. How do I make decisions on the basis of such beliefs, some of which are represented by intervals? There are various rules I could follow for evaluating bets--such as taking the midpoint of each interval and maximizing the mathematical expectation of utility--that would not lead me into violation of the Savage axioms in my betting behavior. If I followed such a rule, I would be ignoring, in my betting behavior, the difference between events whose probability I knew "definitely" (such as Red) and those events whose probabilities were "known" to me only vaguely, within wide intervals. This implication is recognized, and accepted, by Savage, Raiffa and Schlaiffer. As Raiffa and Schlaiffer put it, "We cannot distinguish meaningfully between 'vague' and 'definite' prior distributions." (Applied Statistical Decision Theory, p. 66; their following comment on what they can distinguish is not applicable to this example).

But suppose that I don't follow such a rule. Suppose that I do discriminate, introspectively, between "knowing" and "not knowing" definite probabilities for events, and suppose that I follow a rule in the latter case that leads me to violate the Savage axioms. Specifically, suppose I use the following (quite conservative) decision rule: Act as if one assigned to each event the least favorable of the set of probabilities that might "plausibly" or "reasonably" be associated with it. With the above beliefs, I would act "as if" I assigned $1/6$ probability to Black when I was betting on Black; I would act as if I assigned $1/2$ probability to Black when I was betting against Black (action III; i.e., I would act as if I assigned probability $1/2$ to not-Black). This would generate the observable upper and lower betting probabilities discussed earlier. If I followed this particular decision rule for acting on my intuitive probabilities, the upper and lower betting probabilities would always correspond to my upper intuitive probabilities (if I followed other, plausible rules, my betting probabilities would be enclosed by my intuitive probabilities).

One consequence of using such a rule is that it leads to violations of the Sure-thing Principle when upper and lower intuitive probabilities are unequal. Another is that it biases my choice toward bets with definite probabilities. I will be indifferent between a bet with definite probabilities and one whose probabilities are "ambiguous", only when the "best guess" distribution of pay-offs for the latter is distinctly superior. A third consequence is that if I were offered even odds on Black versus Yellow in the above example (the bet to be off if Red should occur), and I were allowed not to bet, I would prefer not

Professor John W. Pratt

-7-

January 15, 1962
L-1012

to bet rather than to take either end of the bet, even though the payoffs were in utilities. I.e., if I had the option of not betting, I would demand favorable odds on Black before I would bet on Black, and I would also demand favorable odds on Yellow before I would be willing to bet on Yellow. Once again, note that this does not allow you to make a book against me; my acceptable betting odds are "coherent," even though I am violating the Savage axioms.

I should by now have answered your question as to what my own behavior would be, and how I would explain it to myself. What, finally, do I feel about the Sure-thing Principle and the principle of admissibility? I feel that they are normative guides for my own behavior when: (a) the events in question are not "ambiguous," as measured by one of the approaches above; and (b) the payoffs or prizes are not themselves ambiguous bets. If these conditions are not satisfied, I find that certain decision rules--which happen to lead to violations of these two principles in some cases--seem more "reasonable," and more descriptive of my own reflective choices, than any decision rules which would conform strictly to the Savage axioms.

There is a lot more to say, but this seems enough for one letter. My goal, of course, is to convert you, Savage and Raiffa; I'm sending copies of this letter to the latter two. I'm not sure the payoffs are the same as in the Prisoner's Dilemma, but I'm hoping the results will be. Which one will be the first Post-Neo-Bayesian?

Please remember Carol and me to Joy. I'm hoping to be in Cambridge some time this spring, and would like very much to continue this discussion with you and Howard in person.

Yours,

Daniel Ellsberg
Economics Department

DE:esp

Ref. Inc. File No. 35624